

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Caras, Ingrid W
- (ii) TITLE OF INVENTION: A2-1 Neurotrophic Factor
- (iii) NUMBER OF SEQUENCES: 10
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Genentech, Inc.
 - (B) STREET: 1 DNA Way
 - (C) CITY: South San Francisco
 - (D) STATE: California
 - (E) COUNTRY: USA
 - (F) ZIP: 94080
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: WinPatin (Genentech)
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: 08/635130
 - (B) FILING DATE: 19-Mar-1996
 - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Torchia, PhD., Timothy E.
 - (B) REGISTRATION NUMBER: 36,700
 - (C) REFERENCE/DOCKET NUMBER: P1001
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 650/225-8674
 - (B) TELEFAX: 650/952-9881

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1877 base pairs
 - (B) TYPE: Nucleic Acid
 - (C) STRANDEDNESS: Double
 - (D) TOPOLOGY: Linear
- (ix) FEATURE:
 - (A) NAME/KEY: Extra Cellular Domain
 - (B) LOCATION: 244-899
 - (C) IDENTIFICATION METHOD:
 - (D) OTHER INFORMATION:

(ix) FEATURE:

- (A) NAME/KEY: Transmembrane Domain
- (B) LOCATION: 901-978
- (C) IDENTIFICATION METHOD:
- (D) OTHER INFORMATION:

(ix) FEATURE:

- (A) NAME/KEY: signal peptide
- (B) LOCATION: 244-321
- (C) IDENTIFICATION METHOD:
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GNCTAGAAN TAGTGGATCC CCCCAGGCTG CAGGAATTCC GACGGCCCCT 50

GGAAGGGCTC TGGTGGGGCT GAGCGCTCTG CCGCGGGGGC GCGGGCACAG 100

CAGGAAGCAG GTCCCGGTGG GCGCTGGGGG CATCAGCTAC CGGGGTGGTC 150

CGGGCTGAAG AGCCAGGCAG CCAAGGCAGC CACCCCGGGG GGTGGGCGAC 200

TTTGGGGGAG TTGGTGCCCC GCCCCCCCAGG CCTTGGCGGG GTC ATG 246

Met

1

GGG CCC CCC CAT TCT GGG CCG GGG GGC GTG CGA GTC GGG 285

Gly Pro Pro His Ser Gly Pro Gly Val Arg Val Gly

5

10

GCC CTG CTG CTG CTG GGG GTT TTG GGG CTG GTG TCT GGG 324

Ala Leu Leu Leu Leu Gly Val Leu Gly Leu Val Ser Gly

15

20

25

CTC AGC CTG GAG CCT GTC TAC TGG AAC TCG GCG AAT AAG 363

Leu Ser Leu Glu Pro Val Tyr Trp Asn Ser Ala Asn Lys

30

35

40

AGG TTC CAG GCA GAG GGT TAT GTG CTG TAC CCT CAG 402

Arg Phe Gln Ala Glu Gly Gly Tyr Val Leu Tyr Pro Gln

45

50

ATC GGG GAC CGG CTA GAC CTG CTC TGC CCC CGG GCC CGG 441

Ile Gly Asp Arg Leu Asp Leu Leu Cys Pro Arg Ala Arg

55

60

65

CCT CCT GGC CCT CAC TCC TCT CCT AAT TAT GAG TTC TAC 480

Pro Pro Gly Pro His Ser Ser Pro Asn Tyr Glu Phe Tyr

70

75

AAG CTG TAC CTG GTA GGG GGT GCT CAG GGC CGG CGC TGT 519
 Lys Leu Tyr Leu Val Gly Gly Ala Gln Gly Arg Arg Cys
 80 85 90

GAG GCA CCC CCT GCC CCA AAC CTC CTT CTC ACT TGT GAT 558
 Glu Ala Pro Pro Ala Pro Asn Leu Leu Leu Thr Cys Asp
 95 100 105

CGC CCA GAC CTG GAT CTC CGC TTC ACC ATC AAG TTC CAG 597
 Arg Pro Asp Leu Asp Leu Arg Phe Thr Ile Lys Phe Gln
 110 115

GAG TAT AGC CCT AAT CTC TGG GGC CAC GAG TTC CGC TCG 636
 Glu Tyr Ser Pro Asn Leu Trp Gly His Glu Phe Arg Ser
 120 125 130

CAC CAC GAT TAC TAC ATC ATT GCC ACA TCG GAT GGG ACC 675
 His His Asp Tyr Tyr Ile Ile Ala Thr Ser Asp Gly Thr
 135 140

CGG GAG GGC CTG GAG AGC CTG CAG GGA GGT GTG TGC CTA 714
 Arg Glu Gly Leu Glu Ser Leu Gln Gly Val Cys Leu
 145 150 155

ACC AGA GGC ATG AAG GTG CTT CTC CGA GTG GGA CAA AGT 753
 Thr Arg Gly Met Lys Val Leu Leu Arg Val Gly Gln Ser
 160 165 170

CCC CGA GGA GGG GCT GTC CCC CGA AAA CCT GTG TCT GAA 792
 Pro Arg Gly Gly Ala Val Pro Arg Lys Pro Val Ser Glu
 175 180

ATG CCC ATG GAA AGA GAC CGA GGG GCA GCC CAC AGC CTG 831
 Met Pro Met Glu Arg Asp Arg Gly Ala Ala His Ser Leu
 185 190 195

GAG CCT GGG AAG GAG AAC CTG CCA GGT GAC CCC ACC AGC 870
 Glu Pro Gly Lys Glu Asn Leu Pro Gly Asp Pro Thr Ser
 200 205

AAT GCA ACC TCC CGG GGT GCT GAA GGC CCC CTG CCC CCT 909
 Asn Ala Thr Ser Arg Gly Ala Glu Gly Pro Leu Pro Pro
 210 215 220

CCC AGC ATG CCT GCA GTG GCT GGG GCA GCA GGG GGG CTG 948
 Pro Ser Met Pro Ala Val Ala Gly Ala Ala Gly Gly Leu
 225 230 235

GCG CTG CTC TTG CTG GGC GTG GCA GGG GCT GGG GGT GCC 987
 Ala Leu Leu Leu Leu Gly Val Ala Gly Ala Gly Gly Ala

240

245

ATG TGT TGG CGG AGA CGG CGG GCC AAG CCT TCG GAG AGT 1026
 Met Cys Trp Arg Arg Arg Ala Lys Pro Ser Glu Ser
 250 255 260

CGC CAC CCT GGT CCT GGC TCC TTC GGG AGG GGA GGG TCT 1065
 Arg His Pro Gly Pro Gly Ser Phe Gly Arg Gly Gly Ser
 265 270

CTG GGC CTG GGG GGT GGA GGT GGG ATG GGA CCT CGG GAG 1104
 Leu Gly Leu Gly Gly Gly Met Gly Pro Arg Glu
 275 280 285

GCT GAG CCT GGG GAG CTA GGG ATA GCT CTG CGG GGT GGC 1143
 Ala Glu Pro Gly Glu Leu Gly Ile Ala Leu Arg Gly Gly
 290 295 300

GGG GCT GCA GAT CCC CCC TTC TGC CCC CAC TAT GAG AAG 1182
 Gly Ala Ala Asp Pro Pro Phe Cys Pro His Tyr Glu Lys
 305 310

GTG AGT GGT GAC TAT GGG CAT CCT GTG TAT ATC GTG CAG 1221
 Val Ser Gly Asp Tyr Gly His Pro Val Tyr Ile Val Gln
 315 320 325

GAT GGG CCC CCC CAG AGC CCT CCA AAC ATC TAC TAC ACA 1260
 Asp Gly Pro Pro Gln Ser Pro Pro Asn Ile Tyr Tyr Thr
 330 335

TCG ATT TCT GTG TTG GAG TGG CCC ATA TTG CAT ACG ATA 1299
 Ser Ile Ser Val Leu Glu Trp Pro Ile Leu His Thr Ile
 340 345 350

CAA CTG TTT TTC ATG CGA TCC AAG TGC TCC CGT GTC ACT 1338
 Gln Leu Phe Phe Met Arg Ser Lys Cys Ser Arg Val Thr
 355 360 365

ACA TTC TTA TTT CCT GTG CAA GTT ATT ACG ACA TCG ACT 1377
 Thr Phe Leu Phe Pro Val Gln Val Ile Thr Thr Ser Thr
 370 375

TGC CGG ATG ACT TCA TTT AGC TTT ACC ACC CTG AAC CCA 1416
 Cys Arg Met Thr Ser Phe Ser Phe Thr Thr Leu Asn Pro
 380 385 390

TCC ATG CAG GCC TGC AGA GCA CAG ATG GGG GAA TTC CGA 1455
 Ser Met Gln Ala Cys Arg Ala Gln Met Gly Glu Phe Arg
 395 400

ATC AGA TGG TGT TTC TGG GGG GAC AGG ATC CTG GGT ACG 1494
Ile Arg Trp Cys Phe Trp Gly Asp Arg Ile Leu Gly Thr
405 410 415

GCT CTG TTT GTG CTT GTG CTT ATT CTT CTT CTT GGG AGG 1533
Ala Leu Phe Val Leu Val Ile Leu Leu Leu Gly Arg
420 425 430

CTG AAT ATG CAT CAG ACG ACA CTG CTC CGG CAA CGG GCC 1572
Leu Asn Met His Gln Thr Thr Leu Leu Arg Gln Arg Ala
435 440

AGT GTG GAG GCG GAA GCC GGC CAG CAT GGT CCC CTG TG 1610
Ser Val Glu Ala Glu Ala Gly Gln His Gly Pro Leu
445 450 455

ATAGGATTGA AAGAGCTACT GAGAATAGGG GGCTTCTCAA TGAGAGAGCG 1660
GAGGCTGCTG TTATCATGGG AACCAAGGCAG ATCAATCATC CCTGGCAGGT 1710
CAGGCAGGAA GTTACTTAGC TTCTCCTTCA CCTTCTTCCC ACAGAATTAA 1760
TTATAGGCTT GTTCCAAGTT GTAGTGTGTG ATCAGATTG TGCTGCCTGT 1810
CAGCTCTGTG CTACCTGGCA GTTCCCCTCA TGGAAATTGCA TATCAAGCTT 1860
ATCGATACCG TCGACCT 1877

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 455 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Gly Pro Pro His Ser Gly Pro Gly Gly Val Arg Val Gly Ala
1 5 10 15

Leu Leu Leu Leu Gly Val Leu Gly Leu Val Ser Gly Leu Ser Leu
20 25 30

Glu Pro Val Tyr Trp Asn Ser Ala Asn Lys Arg Phe Gln Ala Glu
35 40 45

Gly Gly Tyr Val Leu Tyr Pro Gln Ile Gly Asp Arg Leu Asp Leu
50 55 60

Leu Cys Pro Arg Ala Arg Pro Pro Gly Pro His Ser Ser Pro Asn

65

70

75

Tyr	Glu	Phe	Tyr	Lys	Leu	Tyr	Leu	Val	Gly	Gly	Ala	Gln	Gly	Arg
				80					85					90
Arg	Cys	Glu	Ala	Pro	Pro	Ala	Pro	Asn	Leu	Leu	Leu	Thr	Cys	Asp
				95					100					105
Arg	Pro	Asp	Leu	Asp	Leu	Arg	Phe	Thr	Ile	Lys	Phe	Gln	Glu	Tyr
				110					115					120
Ser	Pro	Asn	Leu	Trp	Gly	His	Glu	Phe	Arg	Ser	His	His	Asp	Tyr
				125					130					135
Tyr	Ile	Ile	Ala	Thr	Ser	Asp	Gly	Thr	Arg	Glu	Gly	Leu	Glu	Ser
				140					145					150
Leu	Gln	Gly	Gly	Val	Cys	Leu	Thr	Arg	Gly	Met	Lys	Val	Leu	Leu
				155					160					165
Arg	Val	Gly	Gln	Ser	Pro	Arg	Gly	Gly	Ala	Val	Pro	Arg	Lys	Pro
				170					175					180
Val	Ser	Glu	Met	Pro	Met	Glu	Arg	Asp	Arg	Gly	Ala	Ala	His	Ser
				185					190					195
Leu	Glu	Pro	Gly	Lys	Glu	Asn	Leu	Pro	Gly	Asp	Pro	Thr	Ser	Asn
				200					205					210
Ala	Thr	Ser	Arg	Gly	Ala	Glu	Gly	Pro	Leu	Pro	Pro	Pro	Ser	Met
				215					220					225
Pro	Ala	Val	Ala	Gly	Ala	Ala	Gly	Gly	Leu	Ala	Leu	Leu	Leu	Leu
				230					235					240
Gly	Val	Ala	Gly	Ala	Gly	Gly	Ala	Met	Cys	Trp	Arg	Arg	Arg	Arg
				245					250					255
Ala	Lys	Pro	Ser	Glu	Ser	Arg	His	Pro	Gly	Pro	Gly	Ser	Phe	Gly
				260					265					270
Arg	Gly	Gly	Ser	Leu	Gly	Leu	Gly	Gly	Gly	Gly	Met	Gly	Pro	
				275					280					285
Arg	Glu	Ala	Glu	Pro	Gly	Glu	Leu	Gly	Ile	Ala	Leu	Arg	Gly	Gly
				290					295					300
Gly	Ala	Ala	Asp	Pro	Pro	Phe	Cys	Pro	His	Tyr	Glu	Lys	Val	Ser
				305					310					315

Gly Asp Tyr Gly His Pro Val Tyr Ile Val Gln Asp Gly Pro Pro
 320 325 330

 Gln Ser Pro Pro Asn Ile Tyr Tyr Thr Ser Ile Ser Val Leu Glu
 335 340 345

 Trp Pro Ile Leu His Thr Ile Gln Leu Phe Phe Met Arg Ser Lys
 350 355 360

 Cys Ser Arg Val Thr Thr Phe Leu Phe Pro Val Gln Val Ile Thr
 365 370 375

 Thr Ser Thr Cys Arg Met Thr Ser Phe Ser Phe Thr Thr Leu Asn
 380 385 390

 Pro Ser Met Gln Ala Cys Arg Ala Gln Met Gly Glu Phe Arg Ile
 395 400 405

 Arg Trp Cys Phe Trp Gly Asp Arg Ile Leu Gly Thr Ala Leu Phe
 410 415 420

 Val Leu Val Leu Ile Leu Leu Gly Arg Leu Asn Met His Gln
 425 430 435

 Thr Thr Leu Leu Arg Gln Arg Ala Ser Val Glu Ala Glu Ala Gly
 440 445 450

 Gln His Gly Pro Leu
 455

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2380 base pairs
 - (B) TYPE: Nucleic Acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GNTCTAGAAN TAGTGGATCC CCCCAGGGCTG CAGGAATTCC GACGGCCCT 50

 GGAAGGGCTC TGGTGGGGCT GAGCGCTCTG CCGCGGGGGC GCGGGCACAG 100

 CAGGAAGCAG GTCCCGGTGG GCGCTGGGGG CATCAGCTAC CGGGGTGGTC 150

 CGGGCTGAAG AGCCAGGCAG CCAAGGCAGC CACCCCGGGG GGTGGGCGAC 200

 TTTGGGGGAG TTGGTGCCCC GCCCCCCCAGG CCTTGGCGGG GTCATGGGGC 250

CCCCCCCATT C TGGGCCGGGG GGC GTGCGAG TCGGGGCCCT GCTGCTGCTG 300
GGGGTTTTGG G GCGTGGTGTC TGGGCTCAGC CTGGAGCCTG TCTACTGGAA 350
CTCGCGAAT AAGAGGTTCC AGGCAGAGGG TGGTTATGTG CTGTACCCTC 400
AGATCGGGGA CCGGCTAGAC CTGCTCTGCC CCCGGGCCCG GCCTCCTGGC 450
CCTCACTCCT CTCCTAATTA TGAGTTCTAC AAGCTGTACC TGGTAGGGGG 500
TGCTCAGGGC CGGCGCTGTG AGGCACCCCC TGCCCCAAAC CTCCTTCTCA 550
CTTGTGATCG CCCAGACCTG GATCTCCGCT TCACCATCAA GTTCCAGGAG 600
TATAGCCCTA ATCTCTGGGG CCACGAGTTC CGCTCGCACC ACGATTACTA 650
CATCATTGCC ACATCGGATG GGACCCGGGA GGGCCTGGAG AGCCTGCAGG 700
GAGGTGTGTG CCTAACCAAGA GGCATGAAGG TGCTTCTCCG AGTGGGACAA 750
AGTCCCCGAG GAGGGGCTGT CCCCCGAAAA CCTGTGTCTG AAATGCCCAT 800
GGAAAGAGAC CGAGGGGCAG CCCACAGCCT GGAGCCTGGG AAGGAGAACCC 850
TGCCAGGTGA CCCCACCAGC AATGCAACCT CCCGGGGTGC TGAAGGCCCC 900
CTGCCCCCTC CCAGCATGCC TGCAGTGGCT GGGGCAGCAG GGGGGCTGGC 950
GCTGCTCTTG CTGGCGTGG CAGGGGCTGG GGGTGCCATG TGTTGGCGGA 1000
GACGGCGGGC CAAGCCTTCG GAGAGTCGCC ACCCTGGTCC TGGCTCCTTC 1050
GGGAGGGGAG GGTCTCTGGG CCTGGGGGGT GGAGGTGGGA TGGGACCTCG 1100
GGAGGCTGAG CCTGGGGAGC TAGGGATAGC TCTGCGGGGT GGCAGGGCTG 1150
CAGATCCCCC CTTCTGCCCT CACTATGAGA AGGTGAGTGG TGACTATGGG 1200
CATCCTGTGT ATATCGTGCA GGATGGGCC CCCCAGAGCC CTCCAAACAT 1250
CTACTACAAG GTATGAGGGC TCCTCTCACG TGGCTATCCT GAATCCAGCC 1300
CTTCTTGGGG TGCTCCTCCA GTTTAATTCC TGGTTTGAGG GACACCTCTA 1350
ACATCTCGGC CCCCTGTGCC CCCCCAGCCC CTTCACTCCT CCCGGCTGCT 1400
GTCCTCGTCT CCACTTTAG GATTCCCTAG GATTCCCCT GCCCCACTTC 1450
CTGCCCTCCC GTTTGGCCAT GGGTGCCCCC CTCTGTCTCA GTGTCCCTGG 1500

ATCCTTTTC CTTGGGGAGG GGCACAGGCT CAGCCTCCTC TCTGACCATG 1550
 ACCCAGGCAT CCTTGTCCCC CTCACCCACC CAGAGCTAGG GGCGGAAACA 1600
 GCCCACCTT TGGTTGGCAC CGCCTCTTT CTGCCTCTCA CTGGTTTCT 1650
 CTTCTCTATC TCTTATTCTT TCCCTCTCTT CCGTCTCTAG GTCTGTTCTT 1700
 CTTCCCTAGC ATCCTCCTCC CCACATCTCC TTTCACCCCTC TTGGCTTCTT 1750
 ATCCTGTGCC TCTCCCATCT CCTGGGTGGG GGCATCAAAG CATTCTCCC 1800
 CTTAGCTTTC AGCCCCCTT CTGACCTCTC ATACCAACCA CTCCCCCTCAG 1850
 TCTGCCAAAA ATGGGGGCCT TATGGGAAG GCTCTGACAC TCCACCCAG 1900
 CTCAGGCCAT GGGCAGCAGG GCTCCATTCT CTGGCCTGGC CCAGGCCTCT 1950
 ACATACTTAC TCCAGCCATT TGGGGTGGTT GGGTCATGAC AGCTACCATG 2000
 AGAAGAAGTG TCCCGTTTG TCCAGTGGCC AATAGCAAGA TATGAACCGG 2050
 TCGGGACATG TATGGACTTG GTCTGATGCT GAATGGGCCA CTTGGGACCG 2100
 GAAGTGACTT GCTCCAGACA AGAGGTGACC AGGCCCGGAC AGAAATGGCC 2150
 TGGGAAGTAG CAGAACAGT GCAGCAGGAA CTGGAAGTGC CTTCATCCAG 2200
 GACAGGAAGT AGCACTTCTG AAACAGGAAG TGGTCTGGCT GGAACTCCAA 2250
 GTGGCTTAGT CTGGGGATC AGGAGGTGGG AGGTGGATGG TTCTTATTCT 2300
 GTGGAGAAGA AGGGCGGGAA GAACTCCTT TCAGGAGGAA GCTGGAACCTT 2350
 ACTGACTGTA AGAGGTTAGA GGTGGACCGA 2380

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 340 amino acids
 - (B) TYPE: Amino Acid
 - (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Gly Pro Pro His Ser Gly Pro Gly Gly Val Arg Val Gly Ala
 1 5 10 15

Leu Leu Leu Leu Gly Val Leu Gly Leu Val Ser Gly Leu Ser Leu
 20 25 30

Glu Pro Val Tyr Trp Asn Ser Ala Asn Lys Arg Phe Gln Ala Glu
 35 40 45

Gly Gly Tyr Val Leu Tyr Pro Gln Ile Gly Asp Arg Leu Asp Leu
 50 55 60

Leu Cys Pro Arg Ala Arg Pro Pro Gly Pro His Ser Ser Pro Asn
 65 70 75

Tyr Glu Phe Tyr Lys Leu Tyr Leu Val Gly Gly Ala Gln Gly Arg
 80 85 90

Arg Cys Glu Ala Pro Pro Ala Pro Asn Leu Leu Leu Thr Cys Asp
 95 100 105

Arg Pro Asp Leu Asp Leu Arg Phe Thr Ile Lys Phe Gln Glu Tyr
 110 115 120

Ser Pro Asn Leu Trp Gly His Glu Phe Arg Ser His His Asp Tyr
 125 130 135

Tyr Ile Ile Ala Thr Ser Asp Gly Thr Arg Glu Gly Leu Glu Ser
 140 145 150

Leu Gln Gly Gly Val Cys Leu Thr Arg Gly Met Lys Val Leu Leu
 155 160 165

Arg Val Gly Gln Ser Pro Arg Gly Gly Ala Val Pro Arg Lys Pro
 170 175 180

Val Ser Glu Met Pro Met Glu Arg Asp Arg Gly Ala Ala His Ser
 185 190 195

Leu Glu Pro Gly Lys Glu Asn Leu Pro Gly Asp Pro Thr Ser Asn
 200 205 210

Ala Thr Ser Arg Gly Ala Glu Gly Pro Leu Pro Pro Pro Ser Met
 215 220 225

Pro Ala Val Ala Gly Ala Ala Gly Gly Leu Ala Leu Leu Leu Leu
 230 235 240

Gly Val Ala Gly Ala Gly Gly Ala Met Cys Trp Arg Arg Arg Arg
 245 250 255

Ala Lys Pro Ser Glu Ser Arg His Pro Gly Pro Gly Ser Phe Gly
 260 265 270

Arg Gly Gly Ser Leu Gly Leu Gly Gly Gly Gly Gly Met Gly Pro
 275 280 285

Arg Glu Ala Glu Pro Gly Glu Leu Gly Ile Ala Leu Arg Gly Gly
290 295 300
Gly Ala Ala Asp Pro Pro Phe Cys Pro His Tyr Glu Lys Val Ser
305 310 315
Gly Asp Tyr Gly His Pro Val Tyr Ile Val Gln Asp Gly Pro Pro
320 325 330
Gln Ser Pro Pro Asn Ile Tyr Tyr Lys Val
335 340

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 454 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GCCGACGCTG TGAGGCACCC CCTGCCCAA ACCTCCTTCT CACTTGTGAT 50
CGCCCAGACC TGGATCTCCG CTTCACCATC AAGTTCCAGG AGTATAGCCC 100
TAATCTCTGG GGCCACGAGT TCCGCTCGCA CCACGATTAC TACATCATTG 150
CCACATCGGA TGGGACCCGG GAGGCCTGGG AGAGCCTGCA GGGAAAGTGTG 200
TGCCTAACCA GAGGCATGAA GGTGCTTCTC CGAGTNGGAC AAAGTCCCGA 250
GGAGGGGCTG TCCCCCGAAA ACCTGTGTCT GAAATGCCA TGAAAGAGA 300
CCGAGGGGCA GCCCACAGCC TGGGAGCCTG GGGAAAGGAGA ACCTGCCAGG 350
TGACCCCACC AGCAATNCAA CCTTCCGGGG TTGCTTGAAG GGCCCCTTGA 400
CCCTTCCCCA GCATTGCNTG CANTTGGTTN GGGGCAGCAN GGGGGNGTTT 450
TGGC 454

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Asp Lys Thr His Thr Cys Pro Pro Cys Pro
1 5 10

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 60 base pairs
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GGACAAAGTC CCGAGGAGGG GCTGTCCCC GAAAACCTGT GTCTGAAATG 50

CCCATGGAAA 60

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 60 base pairs
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CAGGTTCTCC TTCCCCAGGC TCCCAGGCTG TGGGCTGCC CTCGGTCTCT 50

TTCCATGGGC 60

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 346 amino acids
- (B) TYPE: Amino Acid
- (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Met Ala Arg Pro Gly Gln Arg Trp Leu Gly Lys Trp Leu Val Ala
1 5 10 15

Met Val Val Trp Ala Leu Cys Arg Leu Ala Thr Pro Leu Ala Lys
20 25 30

Asn Leu Glu Pro Val Ser Trp Ser Ser Leu Asn Pro Lys Phe Leu
 35 40 45

Ser Gly Lys Gly Leu Val Ile Tyr Pro Lys Ile Gly Asp Lys Leu
 50 55 60

Asp Ile Ile Cys Pro Arg Ala Glu Ala Gly Arg Pro Tyr Glu Tyr
 65 70 75

Tyr Lys Leu Tyr Leu Val Arg Pro Glu Gln Ala Ala Ala Cys Ser
 80 85 90

Thr Val Leu Asp Pro Asn Val Leu Val Thr Cys Asn Arg Pro Glu
 95 100 105

Gln Glu Ile Arg Phe Thr Ile Lys Phe Gln Glu Phe Ser Pro Asn
 110 115 120

Tyr Met Gly Leu Glu Phe Lys Lys His His Asp Tyr Tyr Ile Thr
 125 130 135

Ser Thr Ser Asn Gly Ser Leu Glu Gly Leu Glu Asn Arg Glu Gly
 140 145 150

Gly Val Cys Arg Thr Arg Thr Met Lys Ile Ile Met Lys Val Gly
 155 160 165

Gln Asp Pro Asn Ala Val Thr Pro Glu Gln Leu Thr Thr Ser Arg
 170 175 180

Pro Ser Lys Glu Ala Asp Asn Thr Val Lys Met Ala Thr Gln Ala
 185 190 195

Pro Gly Ser Arg Gly Ser Leu Gly Asp Ser Asp Gly Lys His Glu
 200 205 210

Thr Val Asn Gln Glu Glu Lys Ser Gly Pro Gly Ala Ser Gly Gly
 215 220 225

Ser Ser Gly Asp Pro Asp Gly Phe Phe Asn Ser Lys Val Ala Leu
 230 235 240

Phe Ala Ala Val Gly Ala Gly Cys Val Ile Phe Leu Leu Ile Ile
 245 250 255

Ile Phe Leu Thr Val Leu Leu Lys Leu Arg Lys Arg His Arg
 260 265 270

Lys His Thr Gln Gln Arg Ala Ala Ala Leu Ser Leu Ser Thr Leu
 275 280 285

Ala Ser Pro Lys Gly Gly Ser Gly Thr Ala Gly Thr Glu Pro Ser
 290 295 300
 Asp Ile Ile Ile Pro Leu Arg Thr Thr Glu Asn Asn Tyr Cys Pro
 305 310 315
 His Tyr Glu Lys Val Ser Gly Asp Tyr Gly His Pro Val Tyr Ile
 320 325 330
 Val Gln Glu Met Pro Pro Gln Ser Pro Ala Asn Ile Tyr Tyr Lys
 335 340 345
 Val
 346

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 333 amino acids
- (B) TYPE: Amino Acid
- (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Ala Val Arg Arg Asp Ser Val Trp Lys Tyr Cys Trp Gly Val
 1 5 10 15
 Leu Met Val Leu Cys Arg Thr Ala Ile Ser Lys Ser Ile Val Leu
 20 25 30
 Glu Pro Ile Tyr Trp Asn Ser Ser Asn Ser Lys Phe Leu Pro Gly
 35 40 45
 Gln Gly Leu Val Leu Tyr Pro Gln Ile Gly Asp Lys Leu Asp Ile
 50 55 60
 Ile Cys Pro Lys Val Asp Ser Lys Thr Val Gly Gln Tyr Glu Tyr
 65 70 75
 Tyr Lys Val Tyr Met Val Asp Lys Asp Gln Ala Asp Arg Cys Thr
 80 85 90
 Ile Lys Lys Glu Asn Thr Pro Leu Leu Asn Cys Ala Lys Pro Asp
 95 100 105
 Gln Asp Ile Lys Phe Thr Ile Lys Phe Gln Glu Phe Ser Pro Asn
 110 115 120
 Leu Trp Gly Leu Glu Phe Gln Lys Asn Lys Asp Tyr Tyr Ile Ile
 125 130 135

Ser Thr Ser Asn Gly Ser Leu Glu Gly Leu Asp Asn Gln Glu Gly
 140 145 150
 Gly Val Cys Gln Thr Arg Ala Met Lys Ile Leu Met Lys Val Gly
 155 160 165
 Gln Asp Ala Ser Ser Ala Gly Ser Thr Arg Asn Lys Asp Pro Thr
 170 175 180
 Arg Arg Pro Glu Leu Glu Ala Gly Thr Asn Gly Arg Ser Ser Thr
 185 190 195
 Thr Ser Pro Phe Val Lys Pro Asn Pro Gly Ser Ser Thr Asp Gly
 200 205 210
 Asn Ser Ala Gly His Ser Gly Asn Asn Ile Leu Gly Ser Glu Val
 215 220 225
 Ala Leu Phe Ala Gly Ile Ala Ser Gly Cys Ile Ile Phe Ile Val
 230 235 240
 Ile Ile Ile Thr Leu Val Val Leu Leu Leu Lys Tyr Arg Arg Arg
 245 250 255
 His Arg Lys His Ser Pro Gln His Thr Thr Thr Leu Ser Leu Ser
 260 265 270
 Thr Leu Ala Thr Pro Lys Arg Ser Gly Asn Asn Asn Gly Ser Glu
 275 280 285
 Pro Ser Asp Ile Ile Ile Pro Leu Arg Thr Ala Asp Ser Val Phe
 290 295 300
 Cys Pro His Tyr Glu Lys Val Ser Gly Asp Tyr Gly His Pro Val
 305 310 315
 Tyr Ile Val Gln Glu Met Pro Pro Gln Ser Pro Ala Asn Ile Tyr
 320 325 330
 Tyr Lys Val
 333